

SCIENTIFIC REPORTS



OPEN

The first complete mitochondrial genome for the subfamily Limacodidae and implications for the higher phylogeny of Lepidoptera

Qiu-Ning Liu*, Zhao-Zhe Xin*, Dan-Dan Bian*, Xin-Yue Chai, Chun-Lin Zhou & Bo-Ping Tang

The mitochondrial genome (mitogenome) provides important information for understanding molecular evolution and phylogeny. To determine the systematic status of the family Limacodidae within Lepidoptera, we infer a phylogenetic hypothesis based on the complete mitogenome of *Monema flavescens* (Lepidoptera: Limacodidae). The mitogenome of *M. flavescens* is 15,396 base pairs (bp), and includes 13 protein-coding genes (PCGs), two ribosomal RNA (rRNA) genes, 22 transfer RNA (tRNA) genes, and a control region (CR). The AT skew of this mitogenome is slightly negative and the nucleotide composition is also biased towards A + T nucleotides (80.5%). All PCGs are initiated by ATN codons, except for the cytochrome c oxidase subunit 1 (*cox1*) gene, which is initiated by CGA. All tRNAs display the typical clover-leaf structure characteristic of mitochondrial tRNAs, with the exception of *trnS1* (AGN). The mitogenome CR is 401 bp and consists of several features common to Lepidoptera. Phylogenetic analysis using Bayesian Inference (BI) and Maximum Likelihood (ML) based on nucleotide and amino acid sequences of 13 mitochondrial PCGs indicates that *M. flavescens* belongs to Zygaenoidea. We obtain a well-supported phylogenetic tree consisting of Yponomeutoidea + (Tortricoidea + Zygaenoidea + (Papilionoidea + (Pyraloidea + (Noctuoidea + (Geometroidea + Bombycoidea))))).

The insect mitogenome is a circular molecule 14–19 kilobases in length. It contains 22 tRNAs, 13 PCGs, ATPase subunits 6 and 8 (*atp6* and *atp8*), *cox1-cox3*, *cytochrome B (cob)*, NADH dehydrogenase subunits 1–6 and 4L (*nad1-6* and *nad4L*), the small and large subunit rRNAs (*rrnL* and *rrnS*), and a non-coding element termed the A + T-rich region (CR), which contains initiation sites for transcription and replication^{1,2}. Because of their unique features, including coding content conservation, maternal inheritance, and rapid evolution, mitogenomes have been informative in diverse studies of molecular evolution, such as phylogenetics, population genetics, and comparative and evolutionary genomics^{3,4}.

Recent advances in sequencing technologies have led to the rapid increase in mitogenomic data in Genbank, including Lepidopteran mitogenomes. Lepidoptera is the second largest order of insects, accounting for more than 160,000 species⁵. Zygaenidae is a species-rich superfamily of predominantly diurnal moths with a worldwide distribution. This family is particularly diverse in tropical and subtropical Asia and the Palearctic region⁶. Because of the broad geographical distribution of species, extensive variation in coloration patterns, and an intriguing chemical defence system, Zygaenidae is of great interest to lepidopterists and evolutionary biologists⁷. To date, more than 200 complete or near-complete Lepidopteran mitogenomes are available. However, only one mitogenome of Zygaenoidea has been sequenced⁸. *Monema flavescens* Walker, 1855 is a moth of the Limacodidae family found in Korea, Japan, China, and the Russian Far East. The mitogenome of *M. flavescens* has not been sequenced⁹.

Jiangsu Key Laboratory for Bioresources of Saline Soils, Jiangsu Synthetic Innovation Center for Coastal Bio-agriculture, Jiangsu Provincial Key Laboratory of Coastal Wetland Bioresources and Environmental Protection, School of Ocean and Biological Engineering, Yancheng Teachers University, Yancheng 224051, P. R. China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to Q.-N.L. (email: yctu1989@163.com) or B.-P.T. (email: boptang@163.com)

Superfamily	Family	Subfamily	Species	Genbank No.
Zygaenoidea	Limacodidae		<i>Monema flavescens</i>	KU946971
Zygaenoidea	Zygaenidae	Chalcosiinae	<i>Rhodopsona rubiginosa</i>	KM244668
Bombycoidea	Bombycidae	Bombycinae	<i>Bombyx mandarina</i>	AB070263
Bombycoidea	Bombycidae	Bombycinae	<i>Bombyx mori</i>	AF149768
Bombycoidea	Bombycidae	Bombycinae	<i>Bombyx huttoni</i>	KP216766
Bombycoidea	Bombycidae	Bombycinae	<i>Rondotia menciana</i>	KC881286
Bombycoidea	Saturniidae	Saturniinae	<i>Samia cynthia ricini</i>	JN215366
Bombycoidea	Saturniidae	Saturniinae	<i>Actias selene</i>	JX186589
Bombycoidea	Saturniidae	Saturniinae	<i>Antheraea pernyi</i>	AY242996
Bombycoidea	Saturniidae	Saturniinae	<i>Antheraea yamamai</i>	EU726630
Bombycoidea	Saturniidae	Saturniinae	<i>Eriogyna pyretorum</i>	FJ685653
Bombycoidea	Saturniidae	Saturniinae	<i>Saturnia boisduvalii</i>	EF622227
Bombycoidea	Sphingidae	Sphinginae	<i>Manduca sexta</i>	EU286785
Bombycoidea	Sphingidae	Sphinginae	<i>Sphinx morio</i>	KC470083
Bombycoidea	Saturniidae	Saturniinae	<i>Antheraea frithi</i>	KJ740437
Bombycoidea	Saturniidae	Saturniinae	<i>Attacus atlas</i>	KF006326
Bombycoidea	Saturniidae	Saturniinae	<i>Actias artemis aliena</i>	KF927042
Bombycoidea	Saturniidae	Saturniinae	<i>Samia canningi</i>	KJ159909
Geometroidea	Geometridae	Ennominae	<i>Biston panterinaria</i>	JX406146
Geometroidea	Geometridae	Ennominae	<i>Phthonandria atrilineata</i>	EU569764
Geometroidea	Geometridae	Ennominae	<i>Jankowskia athleta</i>	KR822683
Geometroidea	Geometridae	Ennominae	<i>Apocheima cinerarium</i>	KF836545
Geometroidea	Geometridae	Larentiinae	<i>Operophtera brumata</i>	KP027400
Hepialoidea	Hepialidae		<i>Ahamus yunnanensis</i>	HM744695
Hepialoidea	Hepialidae		<i>Thitarodes renzhiensis</i>	HM744694
Hepialoidea	Hepialidae		<i>Hepialus xiaojinensis</i>	KT834973
Hepialoidea	Hepialidae		<i>Thitarodes pui</i>	KF908880
Hepialoidea	Hepialidae		<i>Thitarodes gonggaensis</i>	KP718817
Hepialoidea	Hepialidae		<i>Napialus hunanensis</i>	KJ632465
Noctuoidea	Erebidae	Lymantriinae	<i>Lymantria dispar</i>	FJ617240
Noctuoidea	Erebidae	Arctiinae	<i>Hyphantria cunea</i>	GU592049
Noctuoidea	Erebidae	Aganainae	<i>Asota plana lacteata</i>	KJ173908
Noctuoidea	Erebidae	Erebinae	<i>Catocala deuteronympha</i>	KJ432280
Noctuoidea	Euteliidae	Euteliinae	<i>Eutelia adulatricoides</i>	KJ185131
Noctuoidea	Nolidae	Chloephorinae	<i>Gabala argentata</i>	KJ410747
Noctuoidea	Nolidae	Risobinae	<i>Risoba prominens</i>	KJ396197
Noctuoidea	Noctuidae	Amphipyridae	<i>Sesamia inferens</i>	JN039362
Noctuoidea	Noctuidae	Amphipyridae	<i>Spodoptera frugiperda</i>	KM362176
Noctuoidea	Noctuidae	Heliothinae	<i>Helicoverpa armigera</i>	GU188273
Noctuoidea	Notodontidae	Phalerinae	<i>Phalera flavescens</i>	JF440342
Noctuoidea	Notodontidae	Thaumetopoeinae	<i>Ochrogaster lunifer</i>	AM946601
Papilionoidea	Hesperiidae	Hesperiinae	<i>Ochlodes venata</i>	HM243593
Papilionoidea	Lycaenidae	Aphnaeinae	<i>Spindasis takanonis</i>	HQ184266
Papilionoidea	Lycaenidae	Theclinae	<i>Protantigius superans</i>	HQ184265
Papilionoidea	Nymphalidae	Apaturinae	<i>Apatura metis</i>	JF801742
Papilionoidea	Nymphalidae	Apaturinae	<i>Sasakia charonda</i>	AP011824
Papilionoidea	Nymphalidae	Calinaginae	<i>Calinaga davidis</i>	HQ658143
Papilionoidea	Nymphalidae	Danainae	<i>Euploea mulciber</i>	HQ378507
Papilionoidea	Nymphalidae	Heliconiinae	<i>Fabriciana nerippe</i>	JF504707
Papilionoidea	Nymphalidae	Libytheinae	<i>Libythea celtis</i>	HQ378508
Papilionoidea	Nymphalidae	Satyrinae	<i>Hipparchia autonoe</i>	GQ868707
Papilionoidea	Papilionidae	Papilioninae	<i>Graphium chironides</i>	KP159289
Papilionoidea	Papilionidae	Papilioninae	<i>Papilio maraho</i>	FJ810212
Papilionoidea	Papilionidae	Parnassiinae	<i>Parnassius apollo</i>	KF746065
Papilionoidea	Papilionidae	Parnassiinae	<i>Parnassius cephalus</i>	KP100655
Papilionoidea	Pieridae	Pierinae	<i>Aporia crataegi</i>	JN796473

Continued

Superfamily	Family	Subfamily	Species	Genbank No.
Papilionoidea	Pieridae	Pierinae	<i>Pieris melete</i>	EU597124
Papilionoidea	Pieridae	Coliadinae	<i>Colias erate</i>	KP715146
Pyraloidea	Crambidae	Spilomelinae	<i>Maruca vitrata</i>	KJ466365
Pyraloidea	Crambidae	Crambinae	<i>Diatraea saccharalis</i>	FJ240227
Pyraloidea	Crambidae	Pyraustinae	<i>Cnaphalocrocis medinalis</i>	JN246082
Pyraloidea	Pyralidae	Galleriinae	<i>Corcyra cephalonica</i>	HQ897685
Pyraloidea	Pyralidae	Phycitinae	<i>Plodia interpunctella</i>	KP729178
Pyraloidea	Pyralidae	Epipaschiinae	<i>Lista haraldusalis</i>	KF709449
Tortricoidea	Tortricidae	Olethreutinae	<i>Rhyacionia leptotubula</i>	JX028270
Tortricoidea	Tortricidae	Olethreutinae	<i>Spilonota lechriaspis</i>	HM204705
Tortricoidea	Tortricidae	Olethreutinae	<i>Grapholita molesta</i>	HQ392511
Tortricoidea	Tortricidae	Tortricinae	<i>Adoxophyes orana</i>	JX872403
Tortricoidea	Tortricidae	Tortricinae	<i>Acleris fimbriana</i>	HQ662522
Tortricoidea	Tortricidae	Tortricinae	<i>Adoxophyes honmai</i>	DQ073916
Yponomeutoidea	Yponomeutidae		<i>Leucoptera malifoliella</i>	JN790955
Yponomeutoidea	Yponomeutidae	Praydinae	<i>Prays oleae</i>	KM874804
Yponomeutoidea	Plutellidae		<i>Plutella xylostella</i>	JF911819

Table 1. List of the complete mitogenomes of Lepidopteran insects.

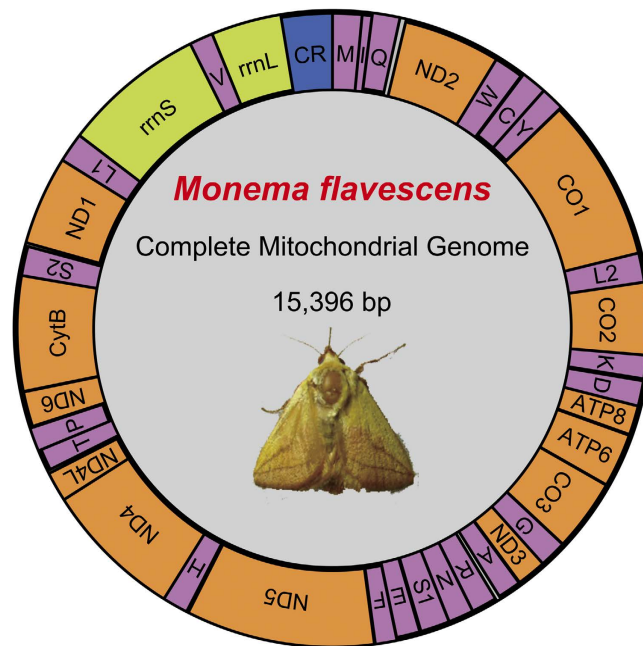


Figure 1. Circular map of the mitogenome of *M. flavescens*. The tRNA genes are labelled according to the IUPAC-IUB. Single-letter amino acids above the bar indicate coding sequence on the major strand, whereas amino acids listed below the bar indicate coding sequence on the minor strand. The one-letter symbols S1, S2, L1, and L2 denote codons *trnS1*(AGN), *trnS2*(UCN), *trnL1*(CUN), and *trnL2*(UUR), respectively.

A better understanding of the Lepidopteran mitogenome requires an expansion of taxon and genome samplings. In this study, we sequence and describe the complete mitogenome of *M. flavescens*. We reconstruct a phylogenetic tree based on PCG sequences in order to analyse the evolutionary relationships among Lepidopteran insects. The assembly and annotation of the *M. flavescens* mitogenome will further the study of Zygaenidea mitochondrial genome architecture and phylogenetics. Furthermore, characterization of the *M. flavescens* mitogenome may provide novel insights into the mechanisms underlying mitogenome evolution.

Methods

DNA Extraction. The moths of *M. flavescens* were collected in Yancheng, Jiangsu Province. Total DNA was isolated using the Genomic DNA Extraction Kit (SangonBiotech, China) according to the manufacturer's instructions. Extracted DNA was used to amplify the complete mitogenome by PCR.

Gene	Direction	Location	Size	Anticodon	Start codon	Stop codon	Intergenic nucleotides
<i>trnM</i>	F	1–68	68	CAT	—	—	2
<i>trnI</i>	F	71–135	65	GAT	—	—	–3
<i>trnQ</i>	R	133–201	69	TTG	—	—	50
<i>nad2</i>	F	252–1266	1015	—	ATT	T	–3
<i>trnW</i>	F	1264–1332	69	TCA	—	—	–9
<i>trnC</i>	R	1324–1388	65	GCA	—	—	5
<i>trnY</i>	R	1394–1462	67	GTA	—	—	3
<i>cox1</i>	F	1466–2996	1531	—	CGA	T	0
<i>trnL2(UUR)</i>	F	2997–3065	69	TAA	—	—	0
<i>cox2</i>	F	3066–3747	682	—	ATG	T	0
<i>trnK</i>	F	3748–3820	73	CTT	—	—	16
<i>trnD</i>	F	3837–3899	63	GTC	—	—	0
<i>atp8</i>	F	3900–4061	162	—	ATT	TAA	–7
<i>atp6</i>	F	4055–4728	674	—	ATG	TAA	3
<i>cox3</i>	F	4732–5517	786	—	ATG	TAA	2
<i>trnG</i>	F	5520–5588	69	TCC	—	—	0
<i>nad3</i>	F	5589–5942	354	—	ATT	TAA	6
<i>trnA</i>	F	5949–6018	70	TGC	—	—	36
<i>trnR</i>	F	6055–6121	67	TCG	—	—	1
<i>trnN</i>	F	6123–6196	74	GTT	—	—	1
<i>trnS1(AGN)</i>	F	6198–6263	66	GCT	—	—	10
<i>trnE</i>	F	6274–6343	70	TTC	—	—	–2
<i>trnF</i>	R	6342–6410	69	GAA	—	—	7
<i>nad5</i>	R	6418–8136	1719	—	ATC	TAA	15
<i>trnH</i>	R	8152–8219	68	GTG	—	—	0
<i>nad4</i>	R	8220–9555	1336	—	ATG	T	0
<i>nad4L</i>	R	9556–9839	284	—	ATA	A	24
<i>trnT</i>	F	9864–9927	64	TGT	—	—	0
<i>trnP</i>	R	9928–9993	66	TGG	—	—	8
<i>nad6</i>	F	10,002–10,520	519	—	ATA	TAA	4
<i>cob</i>	F	10,525–11,674	1150	—	ATG	T	0
<i>trnS2(UCN)</i>	F	11,675–11,744	70	TGA	—	—	–5
<i>nad1</i>	R	11,770–12,708	939	—	ATG	TAA	0
<i>trnL1(CUN)</i>	R	12,709–12,779	71	TAG	—	—	0
<i>rrnL</i>	R	12,780–14,138	1359	—	—	—	0
<i>trnV</i>	R	14,139–14,203	65	TAC	—	—	0
<i>rrnS</i>	R	14,204–14,995	792	—	—	—	0
A + T-rich region		14,996–15,396	401	—	—	—	—

Table 2. Summary of the mitogenome of *M. flavescens*.

PCR Amplification and Sequencing. For amplification of the *M. flavescens* mitogenome, primer sets were designed based upon mitogenomic sequences obtained from other Lepidopteran insects^{10,11}. PCR was performed under the following conditions: 3 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 1–3 min at 48–60 °C, and 10 min at 72 °C. All amplifications were performed on an Eppendorf Mastercycler and Mastercycler gradient in 50 µL reaction volumes. The PCR products were separated by agarose gel electrophoresis (1% w/v) and purified using a DNA gel extraction kit (Transgene, China). The purified PCR products were ligated into the T-vector (Sangon Biotech, China) and sequenced at least three times.

Sequence Assembly and Gene Annotation. Sequence annotation was performed using NCBI BLAST (<http://blast.ncbi.nlm.nih.gov/Blast>) and the DNASTar package (DNASTar Inc. Madison, WI, USA). The identity of tRNA genes was verified using the tRNAscan-SE program (<http://lowelab.ucsc.edu/tRNAscan-SE/>)¹². The nucleotide sequences of PCGs were translated with the invertebrate mitogenome genetic code. Alignments of *M. flavescens* PCGs with various Lepidopteran mitogenomes were performed using Clustal X¹³. Composition skewness was calculated according to the following formulas: AT skew = $[A - T]/[A + T]$; GC ske = $[G - C]/[G + C]$. Codon usage was calculated using MEGA version 6.06. Tandem repeats in the A + T-rich region were predicted using the Tandem Repeats Finder program (<http://tandem.bu.edu/trf/trf.html>)¹⁴.

Phylogenetic Analysis. To reconstruct the phylogenetic relationships among Lepidopteran insects, the complete mitogenomes of Lepidopteran species were obtained from GenBank (Table 1). The amino acid

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	370	1.88	UCU(S2)	99	2.39	UAU(Y)	153	1.82	UGU(C)	29	1.76
UUC(F)	23	0.12	UCC(S2)	14	0.34	UAC(Y)	15	0.18	UGC(C)	4	0.24
UUA(L2)	459	5.1	UCA(S2)	99	2.39	UAA(*)	8	2	UGA(W)	89	1.87
UUG(L2)	25	0.28	UCG(S2)	4	0.1	UAG(*)	0	0	UGG(W)	6	0.13
CUU(L1)	33	0.37	CCU(P)	57	1.81	CAU(H)	62	1.77	CGU(R)	18	1.38
CUC(L1)	0	0	CCC(P)	13	0.41	CAC(H)	8	0.23	CGC(R)	2	0.15
CUA(L1)	23	0.26	CCA(P)	56	1.78	CAA(Q)	60	1.88	CGA(R)	32	2.46
CUG(L1)	0	0	CCG(P)	0	0	CAG(Q)	4	0.13	CGG(R)	0	0
AUU(I)	424	1.86	ACU(T)	76	2.14	AAU(N)	225	1.75	AGU(S1)	18	0.43
AUC(I)	31	0.14	ACC(T)	11	0.31	AAC(N)	32	0.25	AGC(S1)	1	0.02
AUA(M)	238	1.76	ACA(T)	53	1.49	AAA(K)	93	1.69	AGA(S1)	97	2.34
AUG(M)	32	0.24	ACG(T)	2	0.06	AAG(K)	17	0.31	AGG(S1)	0	0
GUU(V)	64	1.74	GCU(A)	72	2.44	GAU(D)	55	1.83	GGU(G)	55	1.11
GUC(V)	0	0	GCC(A)	9	0.31	GAC(D)	5	0.17	GGC(G)	4	0.08
GUA(V)	81	2.2	GCA(A)	34	1.15	GAA(E)	66	1.71	GGA(G)	111	2.23
GUG(V)	2	0.05	GCG(A)	3	0.1	GAG(E)	11	0.29	GGG(G)	29	0.58

Table 3. Codon number and RSCU in *M. flavescens* mitochondrial PCGs. A total of 3,716 codons were analysed, excluding the initiation and termination codons. Amino acids encoded by these codons are labelled according to the IUPAC-IUB single-letter amino acid codes.

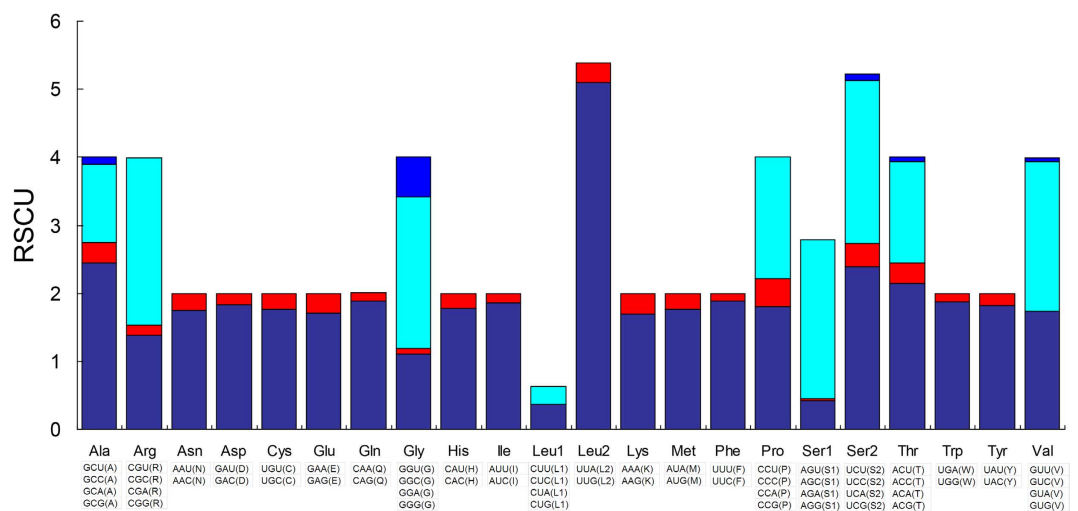


Figure 2. The relative synonymous codon usage (RSCU) in the mitogenome of *M. flavescens*.

<i>M. flavescens</i>	Size (bp)	A%	G%	T%	C%	A + T%	AT skewness	GC skewness
mitogenome	15,396	40.8	7.5	39.7	12.0	80.5	0.014	-0.231
PCGs	11,145	39.9	8.5	38.8	12.8	78.7	0.014	-0.202
tRNAs	1513	41.6	7.5	40.8	10.1	82.4	0.010	-0.148
rRNAs	2151	43.3	4.6	41.2	10.9	84.5	0.025	-0.406
control region	401	43.7	2.00	49.6	4.7	93.3	-0.063	-0.403

Table 4. Composition and skewness in the *M. flavescens* mitogenome.

sequences for each of the 13 mitochondrial PCGs were aligned using default settings and concatenated. This concatenated set of amino acid and nucleotide sequences was used for phylogenetic analysis, which was performed with the Bayesian inference (BI) and Maximum Likelihood (ML) methods using MrBayes v 3.2.2¹⁵ and raxml-GUI, respectively. Alignments of individual genes were performed using MAFFT¹⁶. Gblocks was used to identify conserved regions and remove unreliably aligned sequences within the datasets¹⁷. For the BI and ML analyses, GTR + I + G was the appropriate model for the nucleotide sequences using MrModeltest 2.3 based on Akaike's

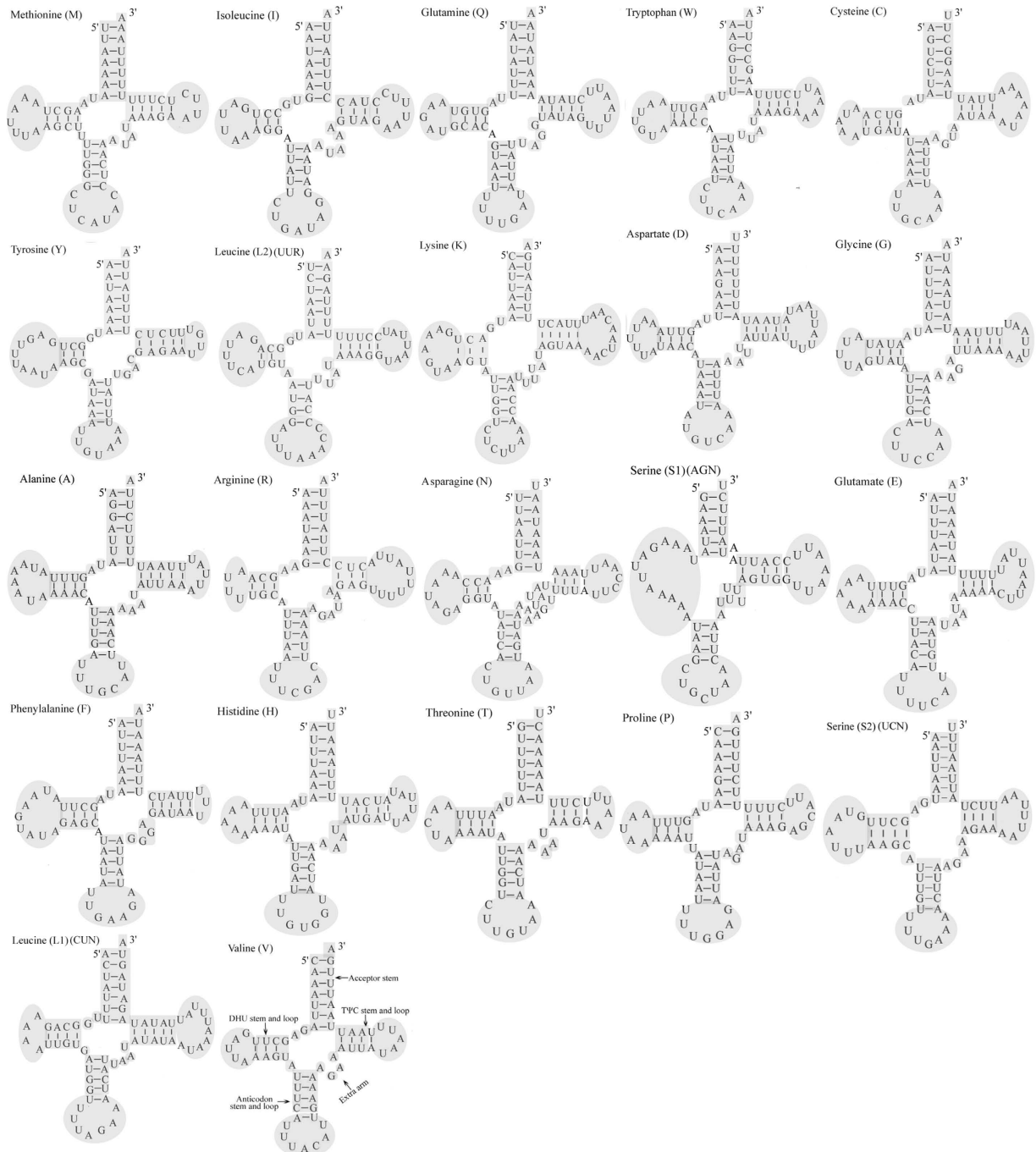


Figure 3. Predicted secondary structures for the tRNA genes in the *M. flavescens* mitogenome.

information criterion (AIC)¹⁸. MtArt + I + G + F was the appropriate model for the amino acid sequence dataset according to ProtTest 3.4 based on AIC¹⁹. Four independent runs were conducted for 10,000,000 generations, and each was sampled every 1,000 generations. All analyses converged within 10,000,000 generations. We assessed the credibility of the results in two ways. First, the average standard deviation of split frequencies was less than 0.05 in the process of Bayesian. Second, we observed sufficient parameter sampling using software Tracer v1.6. The value of ESS was more than 200. This cumulatively suggested that our data was convergent. Posterior probabilities over 0.95 were interpreted as strongly supported. The mitogenomes of Hepialoidea insects were used as outgroups. The resulting phylogenetic trees were visualized in FigTree v1.4.2.

Results and Discussion

Genome Organization and Base Composition. The mitogenome of *M. flavescens* is a closed circular molecule 15,396 bp in size. The gene content is typical of other Lepidopteran insect mitogenomes, including 22 tRNA genes (one for each amino acid and two each for leucine and serine), 13 PCGs (*cox1-3*, *nad1-6*, *nad4L*,

rrnS-14,996ATTATATGTATATATATATA
 ATATAGATTTTTTTTTTTTTTTTTTTTAT
 ATTAATATTTAATATTAATATAATTTT
 TTAATAATTCTTTCTCTCCTAATTT
 ATAATATTTATATTAAAATTAATATCA
 ATATTCCTTCGATATATATTCATTA^{Repeat 1}AAAT
^{Repeat 2}AATTAATATATATATATATTTATAATATA
 TTAAGCTATAAATTATATATTTATATAA
^{Repeat 3}TTAATAATAAATTAAATATTTAATAATA
 TTATTAATTATATAAATAAATTAAATAT
 TTAATATATATATATATATATATACATAA
 TTATGTGTACACATAATTTTTACTTTA
 AAATAAATTAAAAGTTTTTTTTTTTATT
 TATTTTATTATTAAACCGTAATTATAAA
 ATTTACATATAAAAAAAA 15,396-*trnM*

Figure 4. Features present in the A + T-rich region of the *M. flavescens* mitogenome. The reverse strand sequence is shown. Coloured nucleotides indicate the ATATG motif (red), the poly-T stretch (blue), two microsatellite T/A repeat sequences (green), and the poly-A stretch (pink). Two tandem repeats 51 bp in length are indicated in red and black single underline.

cob, *atp6*, and *atp8*), two mitochondrial rRNA genes (*rrnS* and *rrnL*), and a major non-coding region known as the CR. The majority strand (J strand) encodes 23 genes, while the opposite (N) strand encodes 14 genes (Fig. 1, Table 2). The arrangement of the genes within Lepidopteran mitogenomes is usually highly conserved. While the order and orientation of genes in the *M. flavescens* mitogenome are identical to the only other Zygaenoidea insect sequenced to date, this gene order differs from ancestral insects. Specifically, the placement of the *trnM* gene between the CR and *trnI* in the *M. flavescens* mitogenome (CR, *trnM*, *trnI*, *trnQ*, *nad2*) differs from ancestral insects in which *trnM* is located between *trnQ* and *nad2* (CR, *trnI*, *trnQ*, *trnM*, *nad2*)²⁰. However, the ancestral arrangement of the *trnM* gene cluster was also found in ghost moths²¹. This result in *M. flavescens* supports the hypothesis that the ancestral arrangement of the *trnM* gene cluster underwent rearrangement after Hepialoidea diverged from other Lepidopteran lineages. The tRNA gene rearrangements are commonly considered to be a consequence of tandem duplication in a portion of the mitogenome, followed by random or non-random loss of the duplicated copies²².

Skewness, Overlapping, and Intergenic Spacer Regions. The mitogenome of *M. flavescens* has a 29 bp overlap between genes in six locations, with the longest 9 bp overlap located in between *trnW* and *trnC*. The mitogenome of *M. flavescens* contains 167 bp of intergenic spacer sequence spread over 17 regions, ranging in size from 1 to 50 bp (Table 2). The longest spacer sequence is 50 bp located between the *trnQ* and *nad2* genes, and it is extremely A + T rich. The nucleotide composition of the *M. flavescens* mitogenome is as follows: A = 6,275 (40.8%), T = 6,115 (39.7%), G = 1,164 (7.5%), and C = 1,842 (12.0%). As observed in other Lepidopterans, the nucleotide composition of the *M. flavescens* mitogenome is A + T rich (80.5%). This enrichment is lower than in other species, such as *D. punctiferalis* (80.6%), *M. vitrata* (80.7%), *M. testulalis* (80.8%), *L. haraldusalis* (81.5%), and *T. hypsalis* and *N. noctuella* (both 81.4%). In contrast, this enrichment is slightly higher compared to *S. incertulas* (77.1%), *C. suppressalis* (79.7%), and *D. saccharalis* (80.0%). Additionally, the AT skew for the *M. flavescens* mitogenome is slightly positive, indicating a higher occurrence of A compared to T nucleotides. The GC skew values

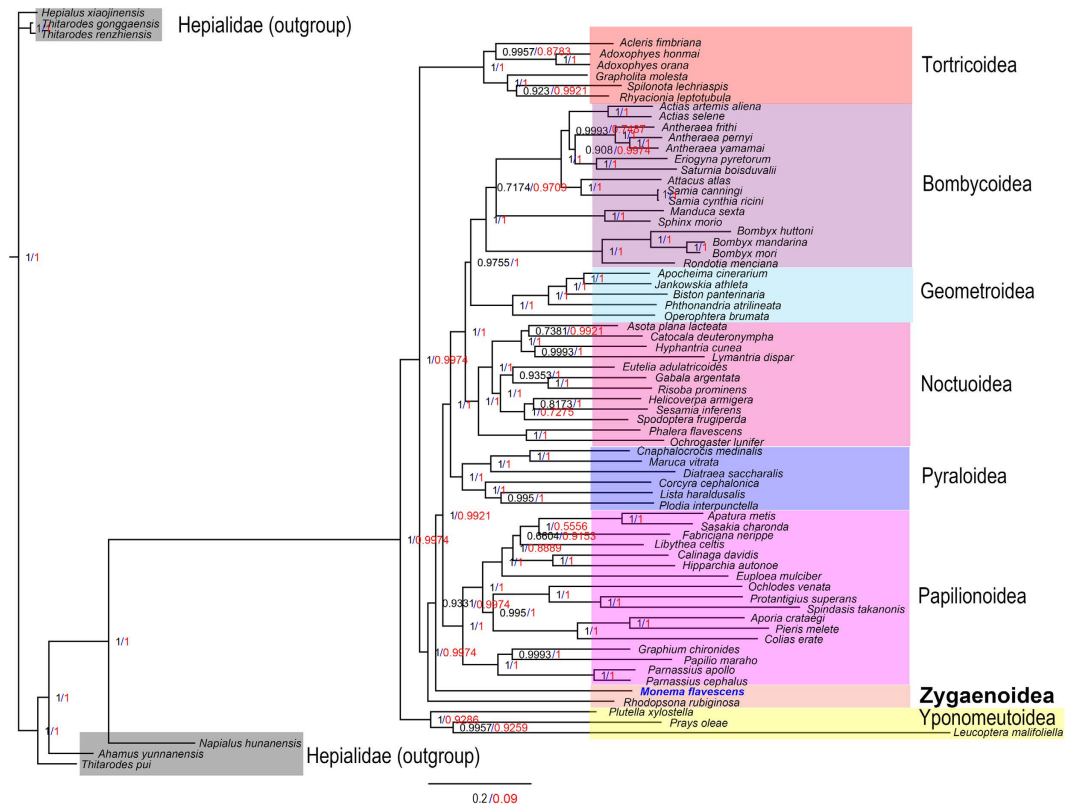


Figure 5. Phylogenetic trees inferred from amino acid (red) and nucleotide (black) sequences of 13 PCGs of the mitogenome using BI analysis.

for the *M. flavescens* mitogenome are negative, indicating a higher content of C compared to G nucleotides. This is similar to GC skew values observed in all sequenced Lepidopteran mitogenomes to date.

Protein-Coding Genes. The start and stop codons of the 13 PCGs in the *M. flavescens* mitogenome are shown in Table 2. Like invertebrate mitogenomes, 12 of these PCGs begin with the standard ATN start codon, except for *cox1*. Sequence alignment revealed that the open reading frame of *cox1* starts with a CGA codon, which encodes arginine. The putative start codon CGA is common in insects^{10,23,24}. An unusual start codon for the *cox1* gene has also been described in various arthropods^{25–27}. In the *M. flavescens* mitogenome, the canonical termination codon, TAA, occurs in seven PCGs. However, the *nad4L* gene utilizes A and the *cox1*, *cox2*, *nad2*, *nad4*, and *cob* genes utilize T as a truncated stop codon instead. Similar results have also been found in other animal mitochondrial genes^{28–31}. Relative synonymous codon usage values for the *M. flavescens* mitogenome are summarized in Table 3 and Fig. 2. The total number of codons in PCGs is 3,716, and the codons CUC, GUC, CCG, UGG, CGG, and AGG are not represented. The most common amino acids in mitochondrial proteins are leucine 2 (Leu 2, 484), isoleucine (Ile, 455), and phenylalanine (Phe, 393), which are likewise highly abundant in mitochondrial proteins in other animals^{32–34}. The average AT content of the 13 PCGs is 78.7%. Furthermore, the AT skew of these PCGs is slightly positive, while the GC skew is slightly negative (Table 4).

Transfer RNA Genes and Ribosomal RNA Genes. The tRNAscan-SE Search Server was used to predict the structure of the 22 tRNAs present in the *M. flavescens* mitogenome. Eight tRNAs are encoded by the L-strand and the remaining 14 are encoded by the H-strand. This tRNA genomic architecture is identical to that found in all Lepidopteran species examined to date. Furthermore, all *M. flavescens* tRNAs display the typical clover-leaf secondary structure observed in most mitochondrial tRNAs with the exception of the *trnS1* (AGN) gene. Interestingly, *trnS1* (AGN) lacking a stable dihydrouridine arm has been observed in several insects, including Lepidopteran species and metazoan mitogenomes^{35–38}. A 7 bp amino acid acceptor stem, in addition to the anticodon stem and loop (7 bp), are both conserved in all tRNAs. While a total of 25 unmatched base pairs were detected in these tRNAs (Fig. 3), 18 of them are G-U pairs, which form a weak bond and are well-known non-canonical pairs in tRNA secondary structures. The remaining seven mismatches include one C-U and six U-U pairs. 22 tRNAs in the *M. flavescens* mitogenome are 1,513 bp long, each of which range in size from 63 to 73 bp. The A + T content is 82.4%. The AT skew for both tRNAs and rRNAs is slightly positive, indicating a higher occurrence of A compared to T nucleotides. The GC skew for both tRNAs and rRNAs is slightly negative, indicating a higher occurrence of C compared to G nucleotides. The two rRNA genes (*rnrS* and *rnrL*) present in *M. flavescens* mitogenome are located between *trnL1* (CUN) and *trnV* or between *trnV* and the A + T-rich region, respectively. The sizes of *rnrL* and *rnrS* are 1,359 bp and 792 bp, respectively. The A + T content of the two rRNAs is 84.5% (Table 4).

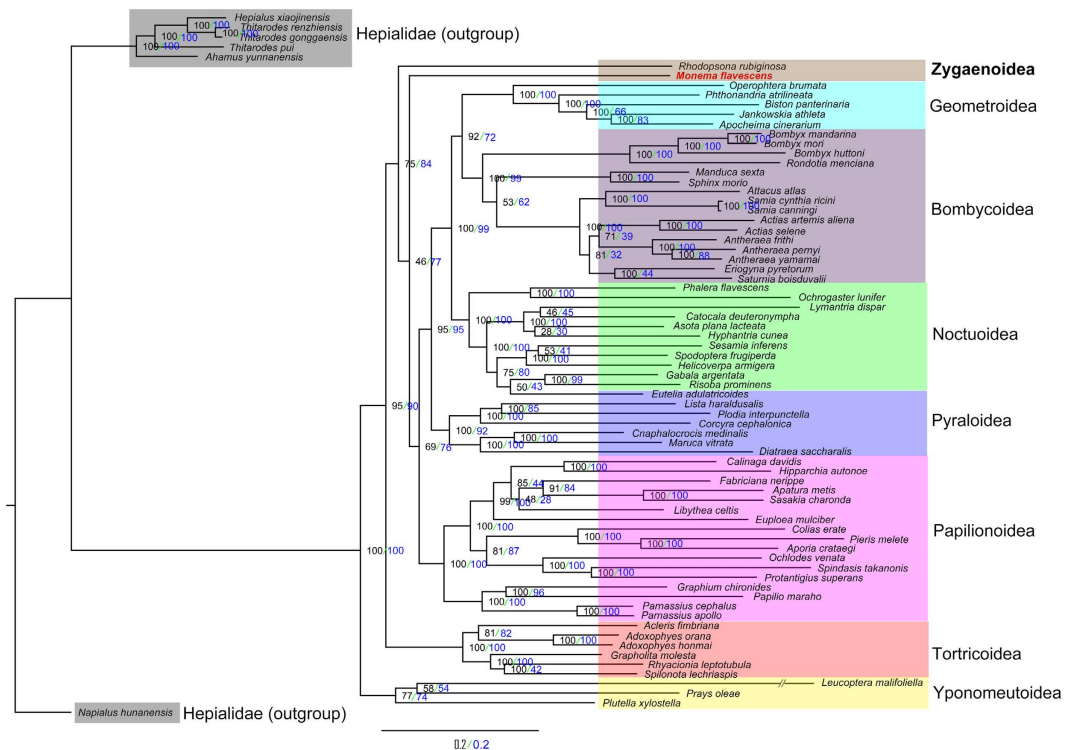


Figure 6. Phylogenetic trees inferred from amino acid (blue) and nucleotide (black) sequences of 13 PCGs of the mitogenome using ML analysis.

Control Region. The CR possesses essential elements involved in the initiation of replication and transcription of the mitogenome³⁹. The CR of the *M. flavescens* mitogenome extends over 401 bp and is located between *rrnS* and *trnM*. The CR contains the highest A + T content (93.3%) in the entire mitogenome. Both the AT skew and GC skew for the CR are slightly negative, indicating that T and C are more abundant than A and G, respectively. Several conserved structures found in other Lepidopteran mitogenomes are also observed in the A + T-rich region of *M. flavescens*. This includes the motif 'ATAGA' and a poly-T stretch downstream of *rrnS*, which is widely conserved in Lepidopteran mitogenomes and may represent the origin of minority or light strand replication⁴⁰. A poly-A commonly observed in other Lepidopteran mitogenomes is also found immediately upstream of the *trnM* gene. We identified microsatellite (AT)₁₀ elements in the A + T-rich region. Multiple tandem repeat elements are typically present in the A + T-rich region of most insects. However, only three tandem repeats are found in the CR of the *M. flavescens* mitogenome (Fig. 4).

Phylogenetic Analyses. Phylogenetic relationships within the Zygaenoidea superfamily are highly debated. In the present study, concatenated amino acid and nucleotide sequences of the 13 PCGs from mitogenomes obtained from nine Lepidopteran superfamilies are used to reconstruct phylogenetic relationships by the BI and ML methods (Figs 5 and 6). The monophyly of each superfamily is generally well supported. The best-supported phylogenetic relationship found in this study is as follows: Yponomeutoidea + (Tortricoidea + Zygaenoidea + (Papilionoidea + (Pyraloidea + (Noctuoidea + (Geometroidea + Bombycoidea))))). The analyses show that *M. flavescens* belongs in the Zygaenoidea superfamily. Both Papilionoidea and Tortricoidea superfamilies are most closely related to Zygaenoidea. More mitogenomes from Zygaenoidea insects were required to resolve the position of Zygaenoidea and the relationships among these superfamilies. Our phylogeny clearly separates and demonstrates a similar topology as that derived from traditional classifications and other molecular data^{41,42}.

References

- Cameron, S. L. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annu. Rev. Entomol.* **59**, 95–117 (2014).
- Wolstenholme, D. R. Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.* **141**, 173–216 (1992).
- Boore, J. L. Animal mitochondrial genomes. *Nucleic Acids Res.* **27**, 1767–1780 (1999).
- Lavrov, D. V. *et al.* Mitochondrial DNA of *Clathrina clathrus* (Calcarea, Calcinea): six linear chromosomes, fragmented rRNAs, tRNA editing, and a novel genetic code. *Mol. Biol. Evol.* **30**, 865–880 (2013).
- Kristensen, N. P., Scoble, M. J. & Karsholt, O. Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfly diversity. *Zootaxa* **1668**, 699–747 (2007).
- Epstein, M. E., Geertsema, H., Naumann, C. M. & Tarmann, G. M. The Zygaenoidea in *Handbook of Zoology* (ed. Kristensen, N. P.) 59–180 (De Gruyter, 1999).
- Naumann, C. M., Tarmann, G. M. & Tremewan, W. G. The Western Palearctic Zygaenidae. *Apollo Books, Stenstrup* (1999).

8. Tang, M. *et al.* Multiplex sequencing of pooled mitochondrial genomes—a crucial step toward biodiversity analysis using mitogenomics. *Nucleic Acids Res.* **42**, e166 (2014).
9. Pan, Z., Zhu, C. & Wu, C. A review of the genus *Monema* Walker in China (Lepidoptera, Limacodidae). *ZooKeys* **306**, 23–36 (2013).
10. Liu, Q. N., Zhu, B. J., Dai, L. S., Wei, G. Q. & Liu, C. L. The complete mitochondrial genome of the wild silkworm moth, *Actias selene*. *Gene* **505**, 291–299 (2012).
11. Liu, Q. N. *et al.* Characterization of the complete mitochondrial genome of the Oriental armyworm, *Mythimna separata* (Lepidoptera: Noctuidae). *Eur. J. Entomol.* **112**, 399–408 (2015).
12. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
13. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882 (1997).
14. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
15. Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
16. Katoh, K. *et al.* MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
17. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).
18. Nylander, J. MrModeltest v2. Program distributed by the author, vol. 2. Evolutionary Biology Centre, Uppsala University (2004).
19. Abascal, F., Zardoya, R. & Posada, D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105 (2005).
20. Boore, J. L., Lavrov, D. V. & Brown, W. Gene translocation links insects and crustaceans. *Nature* **392**, 667–668 (1998).
21. Cao, Y. Q., Ma, C., Chen, J. Y. & Yang, D. R. The complete mitochondrial genomes of two ghost moths, *Thitarodes renzhiensis* and *Thitarodes yunnanensis*: the ancestral gene arrangement in Lepidoptera. *BMC Genomics* **13**, 276 (2012).
22. Juhliling, F. *et al.* Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. *Nucleic Acids Res.* **40**, 2833–2845 (2012).
23. Negrisol, E., Minelli, A. & Valle, G. Extensive gene order rearrangement in the mitochondrial genome of the centipede *Scutigera coleoptrata*. *J. Mol. Evol.* **58**, 413–423 (2004).
24. Zhu, B. J. *et al.* Characterization of the complete mitochondrial genome of *Diaphania pyloalis* (Lepidoptera: Pyralidae). *Gene* **527**, 283–291 (2013).
25. Dai, L. S. *et al.* Characterization of the complete mitochondrial genome of *Cerura menciana* and comparison with other Lepidopteran insects. *PLoS One* **10**, e0132951 (2015).
26. Wang, Y., Liu, X. & Yang, D. The first mitochondrial genome for caddisfly (insecta: Trichoptera) with phylogenetic implications. *Int. J. Biol. Sci.* **10**, 53–63 (2013).
27. Liao, F. *et al.* The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). *Int. J. Biol. Sci.* **6**, 172–186 (2010).
28. Liu, Q. N., Zhu, B. J., Dai, L. S. & Liu, C. L. The complete mitogenome of *Bombyx mori* strain Dazao (Lepidoptera: Bombycidae) and comparison with other lepidopteran insects. *Genomics* **101**, 64–73 (2013).
29. Cameron, S. L. & Whiting, M. F. The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta* (Insecta: lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. *Gene* **408**, 112–113 (2008).
30. Pan, M. H. *et al.* Characterization of mitochondrial genome of Chinese wild mulberry silkworm, *Bombyx mandarina* (Lepidoptera: Bombycidae). *Sci. China C. Life Sci.* **51**, 693–701 (2008).
31. Dai, L. S., Zhu, B. J., Zhao, Y., Zhang, C. F. & Liu, C. L. Comparative mitochondrial genome Analysis of *Eligma narcissus* and other Lepidopteran insects reveals conserved mitochondrial genome organization and phylogenetic relationships. *Sci. Rep.* **6**, 26387 (2016).
32. Song, S. N., Tang, P., Wei, S. J. & Chen, X. X. Comparative and phylogenetic analysis of the mitochondrial genomes in basal hymenopterans. *Sci. Rep.* **6**, 20972 (2016).
33. Liu, Q. N. *et al.* The complete mitochondrial genome of *Eriocheir japonica sinensis* (Decapoda: Varunidae) and its phylogenetic analysis. *Biochem. Syst. Ecol.* **62**, 241–248 (2015).
34. Yong, H. S. *et al.* Complete mitochondrial genome of *Bactrocera arecae* (Insecta: Tephritidae) by next-generation sequencing and molecular phylogeny of Dacini tribe. *Sci. Rep.* **5**, 15155 (2015).
35. Liu, Q. N., Chai, X. Y., Bian, D. D., Zhou, C. L. & Tang, B. P. The complete mitochondrial genome of *Plodia interpunctella* (Lepidoptera: Pyralidae) and comparison with other Pyraloidea insects. *Genome* **59**, 37–49 (2016).
36. Ma, H. *et al.* First mitochondrial genome for the red crab (*Charybdis feriata*) with implication of phylogenomics and population genetics. *Sci. Rep.* **5**, 11524 (2015).
37. Liu, Q. N., Chai, X. Y., Ge, B. M., Zhou, C. L. & Tang, B. P. The complete mitochondrial genome of fall armyworm *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *Genes Genom.* **38**, 205–216 (2016).
38. Li, H. *et al.* Higher-level phylogeny of paraneopteran insects inferred from mitochondrial genome sequences. *Sci. Rep.* **5**, 8527 (2015).
39. Zhang, D. X. & Hewitt, G. M. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. *Biochem. Syst. Ecol.* **25**, 99–120 (1997).
40. Taanman, J. W. The mitochondrial genome: structure, transcription, translation and replication. *Biochim. Biophys. Acta.* **1410**, 103–123 (1999).
41. Kristensen, N. P. & Skalski, A. W. Phylogeny and paleontology. In: Kristensen, N. P. (ed.), *Lepidoptera: Moths and Butterflies, 1 Evolution, Systematics, and Biogeography, Handbook of Zoology Vol IV, Part 35*. Berlin and New York, De Gruyter, 7–25 (1999).
42. Timmermans, M. J., Lees, D. C. & Simonsen, T. J. Towards a mitogenomic phylogeny of Lepidoptera. *Mol. Phylogenet. Evol.* **79**, 169–178 (2014).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (31672267 and 31640074), the Natural Science Foundation of Jiangsu Province (BK20160444), the Natural Science Foundation of Education Department of Jiangsu Province (15KJB240002, 12KJA180009, and 16KJA180008), the Jiangsu Provincial Laboratory of Coastal Wetland Bioresources and Environmental Protection (JLCBE14006), the Jiangsu Key Laboratory for Bioresources of Saline Soils (JKLBS2014013 and JKLBS2015004), the Yancheng City Agricultural Science and Technology Guiding Project (YKN2014022).

Author Contributions

Q.-N.L. and B.-P.T. conceived and designed the experiments. Q.-N.L., Z.-Z.X., D.-D.B. and X.-Y.C., performed the experiments. Q.-N.L. and Z.-Z.X. analyzed the data. C.-L.Z. and B.-P.T. contributed reagents and materials. Q.-N.L. and Z.-Z.X. wrote the paper. Z.-Z.X., Q.-N.L. and B.-P.T. revised the paper.

Additional Information

Accession codes: The *M. flavescens* mitogenome was submitted under the accession number KU946971 to NCBI.

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Liu, Q.-N. *et al.* The first complete mitochondrial genome for the subfamily Limacodidae and implications for the higher phylogeny of Lepidoptera. *Sci. Rep.* **6**, 35878; doi: 10.1038/srep35878 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016